Autograft versus Allograft for Cervical Spinal Fusion: A Systematic Review

Alexander Tuchman1 Darrel S. Brodke2 Jim A. Youssef3 Hans-Jörg Meisel4 Joseph R. Dettori5 Jong-Beom Park6 S. Tim Yoon7 Jeffrey C. Wang8

1 Department of Neurological Surgery, Keck School of Medicine, University of Southern California, Los Angeles, California, United States
2 Department of Orthopedics, University of Utah School of Medicine, Salt Lake City, Utah, United States
3 Durango Orthopedic Associates, P.C./Spine Colorado, Durango, Colorado, United States
4 Department of Neurosurgery, Bergmannstrost Hospital, Halle, Germany
5 Spectrum Research, Inc., Tacoma, Washington, United States
6 Department of Orthopaedic Surgery, Uijeongbu St. Mary’s Hospital, The Catholic University of Korea School of Medicine, Uijeongbu, Korea
7 Department of Orthopedics, Emory Spine Center, Atlanta, Georgia, United States
8 Department of Orthopaedic Surgery, Keck School of Medicine, University of Southern California, Los Angeles, California, United States

Address for correspondence: Alexander Tuchman, MD, Department of Neurological Surgery, Keck School of Medicine, University of Southern California, 1200 N. State Street, Suite 3300, Los Angeles, CA 90033, United States (e-mail: alexandertuchman@gmail.com).

Global Spine J

Abstract

Study Design Systematic review.
Objective To compare the effectiveness and safety between iliac crest bone graft (ICBG), non-ICBG autologous bone, and allograft in cervical spine fusion. To avoid problems at the donor site, various allograft materials have been used as a substitute for autograft. However, there are still questions as to the comparative effectiveness and safety of cadaver allograft compared with autologous ICBG.

Methods A systematic search of multiple major medical reference databases was conducted to identify studies evaluating spinal fusion in patients with cervical degenerative disk disease using ICBG compared with non-ICBG autograft or allograft or non-ICBG autograft compared with allograft in the cervical spine. Radiographic fusion, patient-reported outcomes, and functional outcomes were the primary outcomes of interest. Adverse events were evaluated for safety.

Results The search identified 13 comparative studies that met our inclusion criteria: 2 prospective cohort studies and 11 retrospective cohort studies. Twelve cohort studies compared allograft with ICBG autograft during anterior cervical fusion and demonstrated with a low evidence level of support that there are no differences in fusion percentages, pain scores, or functional results. There was insufficient evidence comparing patients receiving allograft with non-ICBG autograft for fusion, pain, revision, and functional and safety outcomes. No publications directly comparing non-ICBG autograft with ICBG were found.

Conclusion Although the available literature suggests ICBG and allograft may have similar effectiveness in terms of fusion rates, pain scores, and functional outcomes...
following anterior cervical fusion, there are too many limitations in the available literature to draw any significant conclusions. No individual study provided greater than class III evidence, and when evaluating the overall body of literature, no conclusion had better than low evidence support. A prospective randomized trial with adequate sample size to compare fusion rates, efficacy measures, costs, and safety is warranted.

Introduction

Anterior cervical fusion is a safe and effective surgery that continues to experience a rapid increase in utilization. In the United States, cervical fusions increased by 89% from 1993 to 2003, and doubled again between 1998 and 2008. In these cases, surgeons and patients are faced with many decisions, including the choice of bone graft material. Several options are now available, yet autologous iliac crest bone grafting (ICBG) is still considered the gold standard. However, some morbidity is associated with ICBG harvesting that can include infection, hematoma, fracture, wound healing, and donor site pain.

To avoid problems at the iliac crest donor site, other autologous bone has been advocated as a suitable graft material such as local bone and fibular bone. In addition, various allograft materials have been used as a substitute for autologous ICBG. In fact, national trends point to decreased utilization of autograft (86 to 10%) with a reciprocal increase in allograft (14 to 59%) from 1998 to 2004. However, questions remain as to the comparative effectiveness and safety of other types of autograft and cadaver allograft compared with autologous ICBG. Therefore, the purpose of this review was to explore the following key questions:

1. Is autologous ICBG safer and more effective than fusion with other types of autograft in the cervical spine?
2. Is autologous ICBG safer and more effective than cadaver allograft in the cervical spine?
3. Is non-ICBG autograft safer and more effective than cadaver allograft in the cervical spine?

Materials and Methods

Electronic Literature Search

Medline, Embase, and the Cochrane Collaboration Library were systematically searched for literature published through December 21, 2015. The search was limited to studies published in the English language that used human subjects and had abstracts available (see Table 1 in the online supplementary material). Reference lists of key articles from the search as well as applicable systematic reviews were also methodically checked to identify any additional eligible articles. The goal was to identify comparative studies of patients with degenerative joint disease undergoing cervical fusion procedures. Studies using a concurrent control group or a consecutive historical control group (at the same institution) were included, whereas studies with nonconsecutive historical controls or control groups at a different institution were excluded (see Table 2 in the online supplementary material).

Studies were excluded if they did not report results separately by treatment group, used a mixed graft such as demineralized bone matrix, or included skeletal immature patients (< 18 years of age) or patients with a history of tumor or infection in the implantation site, trauma, fracture, or adolescent scoliosis. Studies with a very small sample size (n < 10) for either comparison group were not included. Animal, cadaver, and biomechanical studies were also excluded.

Data Extraction

The following data was extracted: (1) study design, (2) patient characteristics, (3) interventions, (4) inclusion and exclusion criteria, (5) follow-up duration, (6) the rate of follow-up for each treatment group (if reported or calculable), (7) patient-reported functional and pain outcomes (Oswestry Disability Index, visual analog scale, Japanese Orthopaedic Association Score and Scale, Roland-Morris score, Modified MacNab score, or patient satisfaction), (8) various clinical outcomes defined by the investigators, (9) complications or adverse events, (10) fusion rate, (11) time to fusion, (12) definition of fusion, (13) area where bone graft was harvested, (14) type of bone used (i.e., cancellous), (15) preparation methods (i.e., morselization), and (16) preservation method (i.e., freeze-dried or frozen). In the absence of patient-reported or clinical outcomes, radiographic or clinician-defined fusion was used to determine success. Fusion percentages were compared at final follow-up because follow-up times were reported inconsistently across the studies. All extracted data was examined for trends and possible pooling.

Study Quality and Overall Strength

The risk of bias was assessed for each article using criteria set by The Journal of Bone & Joint Surgery, American Volume for therapeutic studies and modified to delineate criteria associated with methodological quality described elsewhere (see Table 3 in the online supplementary material for individual study ratings). After individual article evaluation, the strength of the overall body of evidence with respect to each outcome was determined based on the precepts outlined by the Grades of Recommendation Assessment, Development and Evaluation (GRADE) Working Group and recommendations made by the Agency for Healthcare Research and Quality (AHRQ). Qualitative analysis was performed considering AHRQ-required and additional domains.
<table>
<thead>
<tr>
<th>Key question 2</th>
<th>Table 1 Characteristics of included cervical fusion studies by key question</th>
</tr>
</thead>
</table>
| Inclusion:    | **Autograft versus Allograft for Cervical Spinal Fusion**
<p>| N = 106 (27 or 132 y) | Mean allograft F/U: 4.1 y | N = 106 (27 or 132 y) | Mean allograft F/U: 4.1 y |
| Male:         | NR | Male: | NR |
| N = 53.3%     | NR | N = 53.3% | NR |
| Patient with radicular pain or previous cervical spinal surgery | NR | Patient with radicular pain or previous cervical spinal surgery | NR |
| Inclusion:    | Minimum 12 mo. | Minimum 12 mo. | Minimum 12 mo. |
| Age:          | 46 y | Age: | 46 y |
| N = 23 patients: 94 levels | N = 23 patients: 94 levels | N = 23 patients: 94 levels | N = 23 patients: 94 levels |
| Inclusion:    | Mean 12 mo. | Mean 12 mo. | Mean 12 mo. |
| N = 3 NR      | Mean 12 mo. | Mean 12 mo. | Mean 12 mo. |
| Diagnoses:    | Cervical spondylosis | Cervical spondylosis | Cervical spondylosis |
| N = 3 NR      | None stated | None stated | None stated |
| <em>Autograft</em> | None stated | None stated | None stated |
| <em>Allograft</em> | None stated | None stated | None stated |</p>
<table>
<thead>
<tr>
<th>Author (year); study design; LoE</th>
<th>Intervention/control</th>
<th>Characteristics</th>
<th>Inclusion/exclusion criteria</th>
<th>F/U (range), n/N (% F/U)</th>
<th>Diagnosis</th>
<th>Funding</th>
</tr>
</thead>
</table>
| Parthiban et al (2002)§§; retrospective cohort study; LoE: III | ACDF | N = 68 | Intervention:  
- n = 40  
- Male = NR  
- Age = NR  
- 2 level fusion = 60%  
Control:  
- n = 28  
- Male = NR  
- Age = NR  
- 2 level fusion = 57.1% | Inclusion:  
- NR  
- Exclusion:  
- NR | F/U to 24 mo, % NR | NR | None stated |
| Rish et al (1976)§§; retrospective cohort study; LoE: III | ACDF (Smith-Robinson) | N = 124 | Intervention:  
- n = 80  
- Male = NR  
- Age = NR  
- 2 level fusion = 36.4%  
Control:  
- n = 44  
- Male = NR  
- Age = NR  
- 2 level fusion = 38.7% | (1–12 mo), % NR | Radiculopathy (% NR) | None stated |
| Samartzis et al (2005)§§; retrospective cohort; LoE: III | ACDF with rigid anterior plate fixation | N = 66 | Male = 63.4%  
- Age = 45 (30–83) y  
Intervention:  
- n = 35  
- Male = NR  
- Age = NR  
Control:  
- n = 31  
- Male = NR  
- Age = NR | Inclusion:  
- NR  
- Exclusion:  
- NR | Mean 17 mo (5–60 mo), % NR | Radiculopathy (% NR) | None stated |
| Samartzis et al (2003)§§; retrospective cohort study; LoE: III | ACDF with rigid anterior plate fixation (Smith-Robinson) | N = 77 | Intervention:  
- n = 42 patients, 45 grafts  
- Male = NR  
- Age = 47 y  
- 2 level fusion = 100%  
Control:  
- n = 42 patients, 45 grafts  
- Male = NR  
- Age = 47 y  
- 2 level fusion = 100% | Inclusion:  
- NR  
- Exclusion:  
- NR | Mean 16 mo, % NR | Degenerative spondylolisthesis (% NR) | None stated |
| Young and Rosemwarsser (1993)§§; retrospective cohort, consecutive historical control; LoE: III | ACDF (Smith-Robinson) | N = 48 | Intervention:  
- n = 23  
- Male = 43.5%  
- Age = 35 (23–57) y  
Control:  
- n = 25  
- Male = 52%  
- Age = 37 y | Inclusion:  
- NR  
- Exclusion:  
- NR | NR | Symptomatic cervical disk herniation (% NR) | None stated |
| Zdeblick and Ducker (1991)§§; retrospective cohort; LoE: III | ACDF (Smith-Robinson) | N = 87 | Intervention:  
- n = 27  
- Male = 40.7%  
- Age = 42 range (28–61) y  
- 2 level fusion = 29.6%  
Control:  
- n = 60  
- Male = 45%  
- Age = 44 range (24–68) y  
- 2 level fusion = 31.7% | Inclusion:  
- NR  
- Exclusion:  
- NR | Mean 28 mo (24–41 mo), 87/96 (90.6%) | Disk herniation (49.4%) | None stated |
Table 1 (Continued)

<table>
<thead>
<tr>
<th>Author (year); study design; LoE</th>
<th>Intervention/control</th>
<th>Characteristics</th>
<th>Inclusion/exclusion criteria</th>
<th>F/U (range), n/N (% F/U)</th>
<th>Diagnosis</th>
<th>Funding</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control:</td>
<td>• Fibular strut autograft (structural, 1978–1982)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>n</em> = 59 patients, 118 levels</td>
<td></td>
<td><em>Male = 55.9%</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Age = 46.2 (31–75) y</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 67 patients, 134 levels</td>
<td></td>
<td><em>Male = 55.2%</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Age = 43 y</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inclusion:</td>
<td></td>
<td>• Minimum 24-mo radiographic follow-up after surgery</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Exclusion:</td>
<td></td>
<td>• NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abbreviations: ACDF, anterior cervical diskectomy and fusion; ACF, anterior cervical fusion; F/U, follow up; HNP, herniated nucleus pulposus; ICBG, iliac crest bone graft; LoE, level of evidence; MRI, magnetic resonance imaging; NR, not reported.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aThere was another study group that we were not interested in: VIGOR–r cage (Central Medical Technologies, Taipei, Taiwan) with synthetic, morselized bone.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bDemographic results were only reported for patients available for follow-up.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 Fusion (%) comparing ICBG with allograft in the cervical spine

<table>
<thead>
<tr>
<th>First author and year (study design)</th>
<th>Mean age, y (% male)</th>
<th>F/U, mean (range), mo</th>
<th>Fusion definition</th>
<th>Fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural, freeze-dried allograft</td>
<td></td>
<td></td>
<td>Fusion (via radiograph): bony trabeculae crossing the involved interspace</td>
<td>Allograft</td>
</tr>
<tr>
<td>Bishop 199620 (prospective cohort)</td>
<td>45 (45.5%)</td>
<td>31 (3–43)</td>
<td>Fusion (via radiograph): complete bridging of trabeculae between adjacent vertebral bodies and bone graft</td>
<td>94.3% (50/53)</td>
</tr>
<tr>
<td>Suchomel 200421 (prospective cohort)</td>
<td>47.8 (62%)</td>
<td>39.4</td>
<td>Fusion (via radiograph): no loss of interspace height, no radiolucency across the interspace, good alignment, block configuration, sclerosis of bodies concerned, homogenous amalgamation, no motion of involved joints Poor fusion: anything indicating less than perfect fusion</td>
<td>74% (82/111)</td>
</tr>
<tr>
<td>Rish 197626 (retrospective cohort)</td>
<td>NR (1–12)</td>
<td></td>
<td>Fusion (via radiograph): bony trabeculae clearly seen crossing the disk space</td>
<td>97.8% (21/27)</td>
</tr>
<tr>
<td>Zdeblick 199150 (retrospective cohort)</td>
<td>43 (42.9%)</td>
<td>28 (24–41)</td>
<td>Fusion (via radiograph): bony trabeculae clearly seen crossing the disk space</td>
<td>94.3% (33/35)</td>
</tr>
<tr>
<td>Structural, frozen allografts</td>
<td></td>
<td></td>
<td>Fusion (via radiograph): bony bridge incorporated the graft and adjacent end plates with no radiolucencies or motion</td>
<td>94.3% (33/35)</td>
</tr>
<tr>
<td>Brown 197613 (retrospective cohort)</td>
<td>NR (25.3%)</td>
<td>12</td>
<td>Fusion (via radiograph): complete bridging of trabeculae between adjacent vertebral bodies and the bone graft in less than 20 wk</td>
<td>97.6% (40/41)</td>
</tr>
<tr>
<td>Kao 200524 (retrospective cohort)</td>
<td>55.8 (48.3%)</td>
<td>45</td>
<td>Fusion (via radiograph): no radiolucent line seen on radiograph and no translation or angular change seen on serially lateral F/E radiographs Delayed union: failure of the bone to bridge the interspace and the persistence of a linear lucency seen on radiograph at 6-mo follow-up</td>
<td>90.0% (35/35)</td>
</tr>
<tr>
<td>Samartzis 200328 (retrospective cohort)</td>
<td>48 (NR)</td>
<td>(16–20)</td>
<td>Fusion (via radiograph): bony bridge incorporated the graft and adjacent end plates and when neither instrumentation motion nor radiolucencies were evident encompassing the screws</td>
<td>94.3% (33/35)</td>
</tr>
<tr>
<td>Samartzis 200527 (retrospective cohort)</td>
<td>45 (63.4%)</td>
<td>17 (5–60)</td>
<td>Fusion (via radiograph): presence of a bony bridge incorporating the graft and adjacent end plates and when neither instrumentation motion nor radiolucencies were evident encompassing the screws</td>
<td>100% (35/35)</td>
</tr>
</tbody>
</table>
When the majority of the studies were randomized controlled trials, the initial strength of the overall body of evidence was considered high, and low if otherwise. Published evidence could be downgraded based on the inconsistency of results, indirectness of evidence, imprecision of the effect estimates (e.g., wide confidence intervals), or not having an a priori statement of subgroup analyses. Alternatively, the body of evidence could be upgraded one or two levels if there was a large magnitude of effect or dose–response gradient.

The final overall strength of the body of literature expresses our confidence (high, moderate, or low) that the effect size lies close to the true effect and the extent to which it is believed to be stable based on the adequacy or deficiencies in the body of evidence. A rating of insufficient means that we have very little confidence in the effect estimate; the true effect is likely to be substantially different than the estimated effect. In addition, this rating was used when there was no evidence or it was not possible to estimate an effect.

### Data Analysis
The data was then summarized in tables and further stratified based on the graft tissue preparation and preservation method. The mean differences and variance between baseline and follow-up values were calculated for available continuous variables. Risk proportions (percents) were determined for dichotomous variables by tallying risks as the proportion of patients experiencing an event. When the complication risk was greater in one treatment group compared with another, we calculated the risk ratio and 95% confidence interval using STATA 9.0 (StataCorp., College Station, Texas, United States).

### Results

#### Study Selection
The search strategy yielded 136 potentially relevant citations. Of these, 112 were excluded based on title and/or abstract. Twenty-four were selected for full text review. An additional 11 were excluded based on full text review for the following reasons:...
reasons: not a comparison of interest (n = 4), not a population of interest (n = 3), n < 10 for each group (n = 2), or not a study type of interest (n = 2; see Table 4 in online supplementary material). All 13 included publications evaluated anterior cervical fusion.

Evidence Available

There was no evidence found for key question 1, comparing autologous ICBG with other types of autograft in the cervical spine. The majority of the publications identified supported key question 2, which included two prospective\textsuperscript{20,21} and nine retrospective cohorts.\textsuperscript{22–30} Further details are presented in Table 1.

### Table 4 Complications at final follow-up\textsuperscript{a} comparing allograft versus ICBG in cervical fusion

<table>
<thead>
<tr>
<th>Outcome</th>
<th>First author and year</th>
<th>Graft</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Allograft</td>
</tr>
<tr>
<td>Donor site complications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td>Bishop 1996\textsuperscript{20}</td>
<td>0% (0/32)</td>
</tr>
<tr>
<td></td>
<td>Rish 1976\textsuperscript{26}</td>
<td>0% (0/80)</td>
</tr>
<tr>
<td>Hematoma/seroma</td>
<td>Zdeblick 1991\textsuperscript{30}</td>
<td>0% (0/27)</td>
</tr>
<tr>
<td></td>
<td>Rish 1976\textsuperscript{26}</td>
<td>0% (0/80)</td>
</tr>
<tr>
<td>Wound dehiscence</td>
<td>Rish 1976\textsuperscript{26}</td>
<td>0% (0/80)</td>
</tr>
<tr>
<td>Thigh dysesthesia</td>
<td>Rish 1976\textsuperscript{26}</td>
<td>0% (0/80)</td>
</tr>
<tr>
<td>Osteomyelitis</td>
<td>Rish 1976\textsuperscript{26}</td>
<td>0% (0/80)</td>
</tr>
<tr>
<td>Infection</td>
<td>Bishop 1996\textsuperscript{20}</td>
<td>0% (0/32)</td>
</tr>
<tr>
<td></td>
<td>Rish 1976\textsuperscript{26}</td>
<td>0% (0/80)</td>
</tr>
<tr>
<td>Unsightly scarring</td>
<td>Rish 1976\textsuperscript{26}</td>
<td>0% (0/80)</td>
</tr>
<tr>
<td>Donor site morbidity</td>
<td>Kao 2005\textsuperscript{24}</td>
<td>0% (0/29)</td>
</tr>
<tr>
<td></td>
<td>Chang 2015\textsuperscript{31}</td>
<td>0% (0/44)</td>
</tr>
<tr>
<td>Other complications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superficial wound infection</td>
<td>Young 1993\textsuperscript{28}</td>
<td>4% (1/23)</td>
</tr>
<tr>
<td>Revision</td>
<td>Bishop 1996\textsuperscript{20}</td>
<td>12% (6/50)</td>
</tr>
<tr>
<td></td>
<td>Bose 2001\textsuperscript{22}</td>
<td>25.0% (4/16)</td>
</tr>
<tr>
<td></td>
<td>Zdeblick 1991\textsuperscript{30}</td>
<td>0% (0/27)</td>
</tr>
<tr>
<td>Airway obstruction</td>
<td>Bose 2001\textsuperscript{22}</td>
<td>0% (0/16)</td>
</tr>
<tr>
<td>Deltoid weakness</td>
<td>Bose 2001\textsuperscript{22}</td>
<td>0% (0/16)</td>
</tr>
<tr>
<td>Dysphagia</td>
<td>Bose 2001\textsuperscript{22}</td>
<td>6.3% (1/16)</td>
</tr>
<tr>
<td></td>
<td>Kao 2005\textsuperscript{24}</td>
<td>3.4% (1/29)</td>
</tr>
<tr>
<td></td>
<td>Chang 2015\textsuperscript{31}</td>
<td>2.3% (1/44)</td>
</tr>
<tr>
<td>Laryngeal palsy</td>
<td>Bose 2001\textsuperscript{22}</td>
<td>0% (0/16)</td>
</tr>
<tr>
<td>Hoarseness</td>
<td>Kao 2005\textsuperscript{24}</td>
<td>3.4% (1/29)</td>
</tr>
<tr>
<td>Broken bone graft</td>
<td>Kao 2005\textsuperscript{24}</td>
<td>3.4% (1/29)</td>
</tr>
<tr>
<td>Dislodged bone graft</td>
<td>Kao 2005\textsuperscript{24}</td>
<td>3.4% (1/29)</td>
</tr>
<tr>
<td>Pseudarthrosis</td>
<td>Kao 2005\textsuperscript{24}</td>
<td>0% (0/29)</td>
</tr>
<tr>
<td>Graft settlement or extrusion</td>
<td>Young 1993\textsuperscript{29}</td>
<td>0% (0/27)</td>
</tr>
</tbody>
</table>

Abbreviation: ICBG, iliac crest bone graft.

\textsuperscript{a}See demographics table (Table 2) for final follow up times.

## Efficacy/Effectiveness

### Key Question 1—Iliac Crest Bone Grafting versus non—Iliac Crest Bone Grafting Autograft

There was no evidence found.

### Key Question 2—Iliac Crest Bone Grafting versus Allograft

All but one study showed similar fusion percentages across the groups (Table 2). All together, 74 to 100% of allograft patients were considered fused compared with 62 to 100% of autologous patients. Yet, one prospective study reported a significantly lower proportion of patients with fusion in the allograft group compared with the ICBG group at 3 months.
Autograft versus Allograft for Cervical Spinal Fusion  Tuchman et al.

(50 versus 83%), 12 months (64.0 versus 92%), and final follow-up (88 versus 97.6%, p < 0.05), respectively. There were no apparent differences in fusion percentages using different graft preparation or preservation methods.

Two studies included cervical pain outcomes. One reported no differences in the proportion of patients with mild cervical pain between the allograft group (7%) and ICBG group (5%) at final follow-up. The second reported no difference between groups in the percent improvement in pain from baseline to follow-up: 58.9% in the allograft group compared with 54.1% in the ICBG group.

There were no significant differences noted for outcome measures of clinical success in studies comparing allograft to autologous ICBG for cervical fusion (Odom’s criteria, neck disability index, clinical success; Table 4).

Key Question 3—Non–Iliac Crest Bone Grafting Autograft versus Allograft
In the single publication that addressed key question number 3, the fibular strut autograft group (98/134, 73%) demonstrated a better fusion rate than the freeze-dried fibular allograft group (70/118, 59%; p < 0.01). Mean follow-up was reported to be greater than 24 months. The study did not include axial or radicular pain outcomes.

Safety
Key Question 1—Iliac Crest Bone Grafting versus non–Iliac Crest Bone Grafting Autograft
There was no evidence found.

Key Question 2—Iliac Crest Bone Grafting versus Allograft
Donor site complications among patients who received autologous ICBG varied across studies (Table 4). Donor site pain ranged from 7.2 to 22.7% at final follow-up. Other complications observed at the iliac crest donor site included hematoma/seroma (1.6 to 4.5%), wound dehiscence (6.8%), thigh dysesthesia (4.5%), osteomyelitis (2.3%), infection (2.4 to 6.8%), unsightly scarring (25%), and donor site morbidity (3.8%). Other complications are also detailed in Table 4.

Key Question 3—Non–Iliac Crest Bone Grafting Autograft versus Allograft
Safety outcomes for key question 3 were not reported.

Evidence Summary (Table 5)

Key Question 1—Iliac Crest Bone Grafting versus non–Iliac Crest Bone Grafting Autograft
There was no evidence comparing patients who had received non-ICBG autograft versus ICBG autograft in the cervical spine.

Key Question 2—Iliac Crest Bone Grafting versus Allograft
There was low evidence to support no difference between ICBG and allograft in terms of fusion rates and clinical result.

Though donor site pain and hematoma/seroma occurred more frequently in the ICBG autograft group, there is insufficient evidence to state with reasonable certainty what proportion of patients will experience pain at the donor site due to serious risk of bias in the included studies. There is low evidence around the estimated percent of patients with donor site hematoma/seroma following ICBG harvesting, ranging from 1.6 to 4.5%. There is also low evidence for other donor site complications such as wound dehiscence, thigh dysesthesia, osteomyelitis, unsightly scarring (most frequent), and donor site morbidity occurring more frequently in the ICBG autograft group, ranging from 2.3 to 25% of patients.

Key Question 3—Non–Iliac Crest Bone Grafting Autograft versus Allograft
Only one retrospective cohort study using historical controls that were not applied equally met the inclusion criteria of this review. Thus, the evidence for fusion success, pain, and functional and safety outcomes were deemed insufficient due to imprecision.

Discussion
Based on the current data, during anterior cervical fusion no consistently reported differences in fusion rates or patient outcomes were identified when utilizing autologous ICBG versus allograft. The current state of the published literature is insufficient to comment on the comparative safety of these two graft types or to make any distinctions between ICBG and non-ICBG autograft or non–ICBG autograft and allograft.

Autologous bone graft exhibits many properties that give it theoretical benefits over other materials including its osteoinductive and osteoconductive properties, as well as its inherent growth factors and osteogenic cells that are native to the patient and thus represent lower immune or infection risk. In contradistinction, the overall body of clinical literature failed to show a benefit in terms of long-term fusion rates when comparing ICBG to allograft for anterior cervical fusion. Fusion rates using ICBG or allograft are by far the best-characterized outcome in the literature with 11 studies reporting data on 1,029 patients. But even these publications had limitations in study design that could affect the conclusions. No study had an initial level of evidence rating greater than class III, and all studies had at least one further limitation in relation to lack of blinded, inequality of intervention application, unacceptably high rates of loss to follow-up, inadequate sample size, or lack of control for confounders. Finally, each study had a distinct definition of fusion with no study routinely utilizing computed tomography scan and only five using flexion-extension radiographs.

In interpreting the results of this review regarding fusion rates of allograft versus autograft, it is important to recognize that other patient factors and surgical techniques play a role in the outcome of cervical fusion surgery. Anterior cervical plating has been shown to significantly increase the fusion rate. A meta-analysis by Fraser and Härtl et al found superior fusion rates with anterior cervical plating compared with uninstrumented interbody fusion. Furthermore, there is evidence that the type of...
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Sample size</th>
<th>Risk of bias</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Publication bias</th>
<th>Overall quality of evidence</th>
<th>Treatment groups (%)</th>
<th>Favors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fusion</td>
<td>11 cohorts (n = 1029)</td>
<td>No serious risk of bias</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>No serious risk of imprecision</td>
<td>Undetected</td>
<td>Low</td>
<td>74–100</td>
<td>62–100</td>
</tr>
<tr>
<td>Pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persistent neck pain</td>
<td>1 retrospective cohort (n = 87)</td>
<td>Serious risk of bias</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>No serious risk of imprecision</td>
<td>Undetected</td>
<td>Insufficient due to risk of bias and imprecision</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>% improvement</td>
<td>1 retrospective cohort (n = 82)</td>
<td>Serious risk of bias</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>No serious risk of imprecision</td>
<td>Undetected</td>
<td>Insufficient due to risk of bias and imprecision</td>
<td>58.9</td>
<td>54.1</td>
</tr>
<tr>
<td>Clinical result</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>excellent or good</td>
<td>3 retrospective cohorts (n = 440)</td>
<td>No serious risk of bias</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>No serious risk of imprecision</td>
<td>Undetected</td>
<td>Low</td>
<td>72.4–91.4</td>
<td>68–90.3</td>
</tr>
<tr>
<td>NDI (% improvement)</td>
<td>1 retrospective cohort (n = 82)</td>
<td>Serious risk of bias</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>No serious risk of imprecision</td>
<td>Undetected</td>
<td>Insufficient due to risk of bias and imprecision</td>
<td>67.0</td>
<td>60.8</td>
</tr>
<tr>
<td>Safety</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Donor site pain</td>
<td>2 cohorts (n = 326)</td>
<td>Serious risk of bias</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>No serious risk of imprecision</td>
<td>Undetected</td>
<td>Insufficient due to risk of bias</td>
<td>0</td>
<td>7.2–22.7</td>
</tr>
<tr>
<td>Donor site hematoma/seroma</td>
<td>2 retrospective cohorts (n = 311)</td>
<td>No serious risk of bias</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>No serious risk of imprecision</td>
<td>Undetected</td>
<td>Low</td>
<td>0</td>
<td>1.6–4.5</td>
</tr>
<tr>
<td>Donor site infection</td>
<td>2 cohorts (n = 326)</td>
<td>Serious risk of bias</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>No serious risk of imprecision</td>
<td>Undetected</td>
<td>Insufficient due to risk of bias and imprecision</td>
<td>0</td>
<td>2.4–6.8</td>
</tr>
<tr>
<td>Other donor site complications</td>
<td>2 retrospective cohorts (n = 297)</td>
<td>No serious risk of bias</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>No serious risk of imprecision</td>
<td>Undetected</td>
<td>Low</td>
<td>0</td>
<td>2.3–25</td>
</tr>
<tr>
<td>Infection</td>
<td>1 retrospective cohort (n = 48)</td>
<td>Serious risk of bias</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>No serious risk of imprecision</td>
<td>Undetected</td>
<td>Insufficient due to risk of bias and imprecision</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Revision</td>
<td>3 cohorts (n = 295)</td>
<td>Serious risk of bias</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>No serious risk of imprecision</td>
<td>Undetected</td>
<td>Insufficient due to risk of bias and imprecision</td>
<td>0–25</td>
<td>0–2.2</td>
</tr>
<tr>
<td>Dysphagia</td>
<td>2 retrospective cohorts (n = 179)</td>
<td>Serious risk of bias</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>No serious risk of imprecision</td>
<td>Undetected</td>
<td>Insufficient due to risk of bias and imprecision</td>
<td>3.4–6.3</td>
<td>0–2.2</td>
</tr>
<tr>
<td>Other</td>
<td>3 retrospective cohorts (n = 266)</td>
<td>Serious risk of bias</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>No serious risk of imprecision</td>
<td>Undetected</td>
<td>Insufficient due to risk of bias and imprecision</td>
<td>0–3.4</td>
<td>0–15</td>
</tr>
</tbody>
</table>

Abbreviation: ICBG, iliac crest bone graft.

*a*If evidence was graded insufficient, it was not denoted if one treatment group was favored over the other.

*b*Did not meet two or more important criteria of a good-quality randomized controlled trial or cohort.

*c*Relatively small sample sizes.

*d*Other donor site complications include wound dehiscence, thigh dysesthesia, osteomyelitis, unsightly scarring (most frequent), and donor site morbidity.

*e*Infection indicates deep infection, superficial wound infection, or a non-specified infection.

*f*Other complications include bleeding from segmental vein, damage to lateral femoral cutaneous nerve, postsympathectomy syndrome, airway obstruction, deltoid weakness, laryngeal palsy, hoarseness, broken bone graft, dislodged bone graft, pseudarthrosis, graft settlement, or extrusion.
plating system can affect fusion rates.\textsuperscript{35–38} Only half of the included studies provided information on the use of cervical instrumentation. Four studies used a fixed plating system,\textsuperscript{21,22,27,28} and two studies stated no instrumentation was used.\textsuperscript{20,24} It is important to note that the one study that reported a difference in fusion rates between ICBG and allograft did not use instrumentation,\textsuperscript{20} and the study reporting superiority of fibular strut autograft over allograft did not report on the use of instrumentation.\textsuperscript{32} Other operative factors such as adequate end plate preparation and decortication can also have an effect on achieving fusion.\textsuperscript{39} Patient-related factors such as age, osteoporosis, and tobacco use all seem to have an effect on the fusion rate as well.\textsuperscript{3,20,40}

The main limitation of ICBG for cervical fusion is the remote surgical site required for harvest and the associated morbidity reported to be as high as 22% with up to 90% of patients complaining of donor site pain.\textsuperscript{41,42} Although recent studies have reported significantly lower complication rates especially with allograft reconstruction,\textsuperscript{43–45} Consistent with the overall body of literature, higher rates of donor site complications, such as donor site pain, hematoma/seroma, wound dehiscence, thigh dysesthesias, osteomyelitis, infection, and unsightly scarring, were found in all the pertinent studies related to cervical fusion; however, the overall quality of the evidence related to most donor site complications using the GRADE and AHRQ guidelines was deemed insufficient to draw any conclusion comparing ICBG and allograft. The exceptions were the donor site hematoma or seroma group and the “other” donor site complications groups where there was low evidence of no statistically significant difference between the two interventions. This result likely represents a type II error as few studies reported complications and those that did tended to have low morbidity rates leading to a small sample available for analysis.

Though using an allograft removes the donor site risks from the patient while maintaining osteoconductive properties, it tends to incorporate more slowly, leading to longer times to fusion, and is associated with higher direct costs.\textsuperscript{4} Furthermore, there are concerns that immunologic mismatch or inflammatory reactions to the products used to prepare and sterilize the allograft can put a patient at increased risk for complications.\textsuperscript{3,4,20} The preparation process required of the allograft prior to implantation has been reported to weaken it,\textsuperscript{36–49} hinting at a mechanism for loss of correction or even revision surgery. It should also be noted that there have been extremely rare infection risks associated with allograft including 2 cases of human immunodeficiency virus, 3 cases of viral hepatitis, and 26 bacterial infections.\textsuperscript{50–53}

This study highlights significant limitations in the available literature and multiple avenues for further study with respect to graft choice for anterior cervical fusion. There is insufficient evidence to compare the safety and efficacy of cervical fusion when using ICBG versus non-ICBG autograft (key question 1) and non-ICBG autograft versus allograft (key question 3). Although 11 total cohorts directly compared cervical fusion with ICBG autograft versus allograft, no studies were randomized and only two studies collected data prospectively. Furthermore, there was little consistency between studies with regards to follow-up time, definition of fusion, and reporting of outcomes and complications. This inconsistency significantly limited our ability to pool the data and perform a formal meta-analysis. Thus the overall quality of the existing literature comparing these two graft types remains limited. Future studies on this topic should be aimed at comparing efficacy measures, direct and indirect costs, and safety in a prospective fashion. Future study design must include sufficient power to assess clinically relevant complications, adequate long-term follow-up, and a reliable and reproducible method to define fusion.

Conclusion

The best available evidence weakly suggests that ICBG and allograft demonstrated similar effectiveness in terms of fusion rates, pain scores, and functional outcomes following anterior cervical fusion. However, significant limitations in the available literature were obvious. Therefore, definitive judgments or suggestions with respect to the use of ICBG or allograft should be made carefully and within the framework of the current literature. At this time, ICBG versus other fusion methods remains an area of clinical equipoise and thus is not only an interesting area for further investigation but necessary. A well-designed randomized controlled trial comparing ICBG and allograft for anterior cervical fusion is warranted to address questions related to differential fusion rates, clinical efficacy, cost, and safety.

Disclosures

Alexander Tuchman: Grant (NuVasive)
Darrel S. Brodke: Royalties (Amedica, Depuy Synthes, Medtronic)
Jim A. Yousef: Consultant (NuVasive, Integra, Amedica, HealthTrust); Royalties (NuVasive, Osprey Medical, Integra, Amedica)
Hans-Joerg Meisel: Consultant (Zyga, Diffusion, Codon); Royalties (Medtronic, Aesculap, Fehling)
Jong-Beom Park: none
S. Tim Yoon: Consultant (ISSLS, Stryker, Meditech); Grant (Biomet, NIH, Pfizer); Royalties (Stryker, Biomet, Meditech); Stock ownership (Alphatec, Phygen, Medyssey, Meditech)
Jeffery C. Wang: Personal fees (AO Foundation, NASS, CSRS, CSRF); Royalties (Biomet, Stryker, Alphatec, Synthes, Amedica, Osprey, Aesculap, Seapine); Stock ownership (Fziomed, Prometheus Spine, Paradigm Spine, Benevenue, NexGen, Pioneer, Amedica, Vertiflex, Electrocore, Surgitech, Axiomed, VD Innovations, CoreSpine, Expanding Orthopaedics, Syndicom, Osprey, Bone Biologics, Curative Biosciences, PearlDiver, Pioneer)

Acknowledgments

Analytic support for this work was provided by Spectrum Research, Inc. with funding from the AOSpine Foundation.
References

40 DeWald CJ, Stanley T. Instrumentation-related complications of multilevel fusions for adult spinal deformity patients over age 65.
44 Devine JG. Bone grafting techniques in idiopathic scoliosis: a confirmation that allograft is as good as autograft but dispels the purported pain associated with the iliac crest bone graft harvest. Spine J 2013;13(5):530–531